Privacy and Identity Management in Europe for Life

Towards Usable Privacy Enhancing Technologies: Lessons Learned from the PrimeLife Project

Editors:	Cornelia Graf (CURE)
	Christina Hochleitner (CURE)
	Peter Wolkerstorfer (CURE)
	Julio Angulo (KAU)
	Simone Fischer-Hübner (KAU)
	Erik Wästlund (KAU)
Reviewers:	Jan Camenisch (IBM)
	John Sören Pettersson (KAU)
	Ronald Leenes (TILT)

Identifier:	D4.1.6
Type:	Deliverable
Class:	Public
Date:	June 17, 2011

Abstract

In this deliverable, we present lessons learnt from the PrimeLife HCI (Human Computer Interaction) Activity by discussing typical HCI challenges and fallacies that we experienced during the PrimeLife project. We also provide guidance on how these issues can be addressed in order to develop usable privacy-enhancing technology solutions.

The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement n° 216483 for the project PrimeLife.

Members of the PrimeLife Consortium

1.	IBM Research GmbH	IBM	Switzerland
2.	Unabhängiges Landeszentrum für Datenschutz	ULD	Germany
3.	Technische Universität Dresden	TUD	Germany
4.	Karlstads Universitet	KAU	Sweden
5.	Università degli Studi di Milano	UNIMI	Italy
6.	Johann Wolfgang Goethe – Universität Frankfurt am Main	GUF	Germany
7.	Stichting Katholieke Universiteit Brabant	TILT	Netherlands
8.	GEIE ERCIM	W3C	France
9.	Katholieke Universiteit Leuven	K.U.Leuven	Belgium
10.	Università degli Studi di Bergamo	UNIBG	Italy
11.	Giesecke & Devrient GmbH	GD	Germany
12.	Center for Usability Research & Engineering	CURE	Austria
13.	Europäisches Microsoft Innovations Center GmbH	EMIC	Germany
14.	SAP AG	SAP	Germany
15.	Brown University	UBR	USA

List of Contributors

This deliverable has been jointly authored by multiple PrimeLife partner organisations. The following list presents the contributors for the individual parts of this deliverable.

Chapter	Author(s)
Executive Summary	Simone Fischer-Hübner (KAU)
Introduction	Simone Fischer-Hübner (KAU)
HCI Fallacies to be considered during UI Design and Testing	Julio Angulo (KAU), Simone Fischer-Hübner (KAU), Cornelia Graf (CURE), Erik Wästlund (KAU), Peter Wolkerstorfer (CURE), Erik Wästlund (KAU)
How to design PET User Interfaces	Julio Angulo (KAU), Simone Fischer-Hübner (KAU), Cornelia Graf (CURE), Christina Hochleitner (CURE), Peter Wolkerstorfer (CURE)
How to evaluate PET User Interfaces	Julio Angulo (KAU), Simone Fischer-Hübner (KAU), Cornelia Graf (CURE), Christina Hochleitner (CURE), Erik Wästlund (KAU), Peter Wolkerstorfer (CURE)
Conclusions	Simone Fischer-Hübner (KAU)

Executive Summary

The development of usable privacy-enhancing Identity Management poses several HCI (Human Computer Interaction) challenges. This deliverable reports about typical HCI fallacies and challenges that we and other HCI researcher experienced when developing user interfaces for Privacy-Enhancing Technologies (PETs). Furthermore, it reports about lessons that we learned from the PrimeLife HCI Activity work and provides guidelines for the design of usable PETs.

After the background and structure of this deliverable has been presented in Chapter 1 ("Introduction"), Chapter 2 ("HCI Fallacies to be considered during UI Design and Testing") starts with the discussion of some of the major HCI fallacies that we experienced and that should be considered during the UI (User Interface) design and testing. This includes the problem of many users to differentiate whether data is stored on the user side (under the user's control) or on a remote services side and the problem to comprehend to which network entities personal data flows during online transactions. User interfaces have therefore to address the challenge to evoke the correct mental model in regard to where data are transferred to and where they are processed. Another issue is that privacy warnings can cause rushed and unwanted user reactions and thus needs to be designed with care. Furthermore, the mediation of trustworthiness, intercultural differences and a well comprehensible terminology to be used in UIs are challenges to be taken into consideration. Many of the HCI issues that we experienced are mental model issues which are difficult to solve for novel PET concept, which are unfamiliar for the users. This is especially true for those PETs, for which no obvious real world analogies exist, such as for instance for anonymous credentials and their selective disclosure properties. This shows once more that evoking adequate mental models is a key issue for the successful deployment of novel privacy technologies.

Based on our experiences and the lessons learned, and based on the research results of others, Chapter 3 ("How to design PET User Interfaces") provides HCI guidelines for the design of usable PET user interfaces. For this, HCI heuristics for PETs are provided, which adapt, extend and exemplify the classical list of Nielsen's Usability Heuristics for the PET domain. Besides, a brief overview to the HCI Patterns for PETs that we developed in PrimeLife is given, which provide best practice solutions for the PET user interface design and which should be applied in combination with the Usability Heuristics. The chapter concludes with discussing the need for following a User-Centric Design Process for developing PETs and showing how HCI patterns and heuristics should be applied during the User-Centric Design Cycles.

In Chapter 4 ("How to evaluate PET User Interfaces"), we present important factors for the planning and performing of usability evaluations of privacy-enhancing technologies, which are based on our experience in executing usability evaluations for PrimeLife project prototypes. In particular, when recruiting test participants, aspects such as their cultural and technical background need to be considered. The wording and privacy terminology used in the evaluations needs to be carefully chosen, and test tasks need to be designed with care for evaluating the user's understanding of the PET functionality. Participants need to be introduced to the tests in a consistent and comprehensible manner and demographic data that are valuable for later analysis in particular in regard to the test user's technical and cultural backgrounds need to be collected. Factors that might have an influence on the participants' perception of privacy risks need to be considered when carrying out the evaluation, and finally, post test questionnaires and PET-USES (Privacy-Enhancing Technology Users' Self-Estimation Scale), which was developed in PrimeLife, should be used for obtaining a more accurate account of the experience and opinions of the test participants.

As Chapter 5 (“Conclusions”) concludes, this deliverable provides an experience report, which can help UI developers for PET solutions to avoid doing typical mistakes and provides at the same time HCI heuristics, best practice solutions and guidance for the development of usable PETs.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Privacy-enhancing Identity Management will only be successful if its technologies are accepted and applied by end users. For this reason, PrimeLife Activity 4 has had the objective to research and develop user interfaces for PrimeLife technologies, which are intelligible, trustworthy and user-friendly while being compliant with legal privacy principles. To accomplish these objectives, several challenges had to be met – many of them were related to the difficulty of inducing the correct mental models for novel PETs (privacy-enhancing technologies). Also, the fact that security and privacy-related decisions often appear in a context, when they are only of secondary interest for users who are rather focussed on getting their primary task completed (e.g., purchase order), raises special challenges for the design of user interfaces (UIs) and evaluation of PETs.

The objective of this deliverable is to present our experiences and lessons learnt from our HCI work in the PRIME[1] and PrimeLife projects, and, based on these findings as well as other research results, to provide guidelines and best practice methods for the design and evaluation of usable PETs. It therefore extracts those key findings from the other PrimeLife HCI (Human Computer Interaction) deliverables[2], which provide advice to others that are developing and/or designing user interfaces for PETs.

The remainder of this deliverable is structured as follows:

In Chapter 2 ("HCI Fallacies to be considered during UI Design and Testing for PETs"), we discuss some major HCI problems and fallacies that we experienced in the PRIME and PrimeLife projects. For each of those fallacies, we summarise the lessons that we learned in terms of what needs to be considered during UI development and evaluation in order to address those fallacies.

Chapter 3 ("How to design PET User Interfaces") provides guidance on how to design user interfaces for privacy-enhancing technologies. For this, it starts by presenting HCI heuristics, which are derived from our lessons learned and from other research results. Then, we provide an overview to HCI patterns that we have developed in PrimeLife, which are merging best practice solutions from HCI and different guidelines, including those that we derived for PrimeLife as an

[1] EU PF6 project PRIME (Privacy and Identity Management for Europe), https://www.prime-project.eu/

[2] Other PrimeLife HCI deliverables can be found at: http://www.primelife.eu/results/documents

approach to describe, organize und present solutions and best practices for design problems in the PET domain. Finally, the need for a User-Centred Design Process for developing PETs is discussed and it is shown how both HCI heuristics and HCI patterns can be applied during the User-Centric Design Process.

Chapter 4 (“How to evaluate PET User Interfaces”) continues by presenting advice on how to evaluate user interfaces for privacy-enhancing technologies, and what needs to be considered for the preparation and performance of the usability tests.

Finally, Chapter 5 (“Conclusions”) will round up this deliverable by drawing some overall conclusions.

Chapter 2

HCI Fallacies to be considered during UI Design and Testing for PETs

Throughout our HCI research and development work that we conducted in the PRIME and PrimeLife projects, we spotted several typical usability problems and HCI fallacies that occur with PETs and privacy-enhancing identity management in particular.

In this chapter, we will describe some of the major HCI fallacies that we experienced. We also refer to some related work, as far as it is inline with our experiences. For each fallacy we summarise our lessons learned and especially point out what needs to be considered during the UI design and testing of PETs in order to approach these fallacies. Finally, at the end of this chapter, we underline the importance of inducing adequate mental models for PETs.

2.1 Users' Assumptions of data handling and data flow on the Internet

For understanding privacy implications and for making well-informed decisions in regard to the disclosure of one's personal data, users should understand who actually receives their data and the way it is processed. In its simplest form, an online transaction involves a user sending data and a service provider receiving that data. However, this is rarely the case since, for example, a shopping transaction might include a payment service and a delivery service in addition to the vendor, who also have access to the users' data. To further complicate the issue, users' data can be collected by third party trackers in order to build profiles of users to make revenue. Not only are users often not aware of this data collection taking place, but they also have no information to which parties their data flow and where their data are stored.

Several prototypes for PETs have been developed within the PRIME and PrimeLife projects that have shown the difficulty of users developing appropriate mental models for these technologies. For instance, a prototype named the Data Track was developed to promote the concept of transparency and to give users the possibility to learn which data they have released in the past, where they are stored, and to access and correct these data online if they wish to do so. Usability studies of this prototype showed that users have difficulties differentiating the information being handled at the user-side and the one handled at the services-side [Pri422]; a difficulty that was also

already recognized in the PRIME project [PFHD+05]. Other prototypes, such as early versions of the "Send Data?" dialog [Pri432] and of anonymous credential selectors confirmed that participants found it hard to recognize how their data were transferred to different entities. Results showed that some participants believed that, at the moment of an online shopping transaction, their requested information was sent to the issuer of a certificate as well as to the service provider in question. For example, users believed that if the police had issue a passport certificate, then their information would also go through the police at the moment of the transaction [Pri414].

In regards to data storage, the movement towards constant internet access and cloud computing has blurred the line between user-side and services-side data management. In the middle of the nineties, when internet was accessed via dial-up modems this dichotomy was very clear. Users would dial up the modem pool, download e-mails and then explicitly disconnect from the internet by hanging up. Now, with broadband connections in combination with data that are stored on-line and accessible through web interfaces or locally running software alike it is no longer as explicit were data are stored, neither is there a clear distinction on whether a computing device is connected to the Internet or not, and when it is transmitting data. Taken together, the voluntary release of personal data, the involuntary release of behavioral data, and the cloud based pooling of resources add to the complexity for users to create comprehensive mental models of who receives, collects, stores, and uses their personal data. Future computer paradigms involving cloud computing and networked devices embedded in everyday objects bring even further challenges to the mental models of information flow.

2.1.1.1.1 Lessons learned:

Users often do not have a correct understanding of where (at what site) their personal data is stored and processed and to what entities their data is transferred. When designing and testing privacy-enhancing identity management systems, investigations are thus needed on how to evoke the correct mental models in users with regard to where what data are transmitted and under whose control the data are stored and processed. Having a comprehensive mental model will be essential for them to estimate privacy risks correctly, to understand better how far PETs can protect their online privacy.

HCI techniques need to be used for prominently illustrating whether the user- or the services side is concerned. For example, a trust evaluation function developed in PrimeLife [Pri421] clarified through wording and user interface structure that the trustworthiness of a contacted services side (and not the trustworthiness of the user's computer) was evaluated. For the "Send Data?" user interface developed for the PrimeLife Policy Engine, the service provider's website behind the "Send Data?" is dimmed, helping users understand that the dialog works on the client side and is not part of the service providers (see Figure 1)

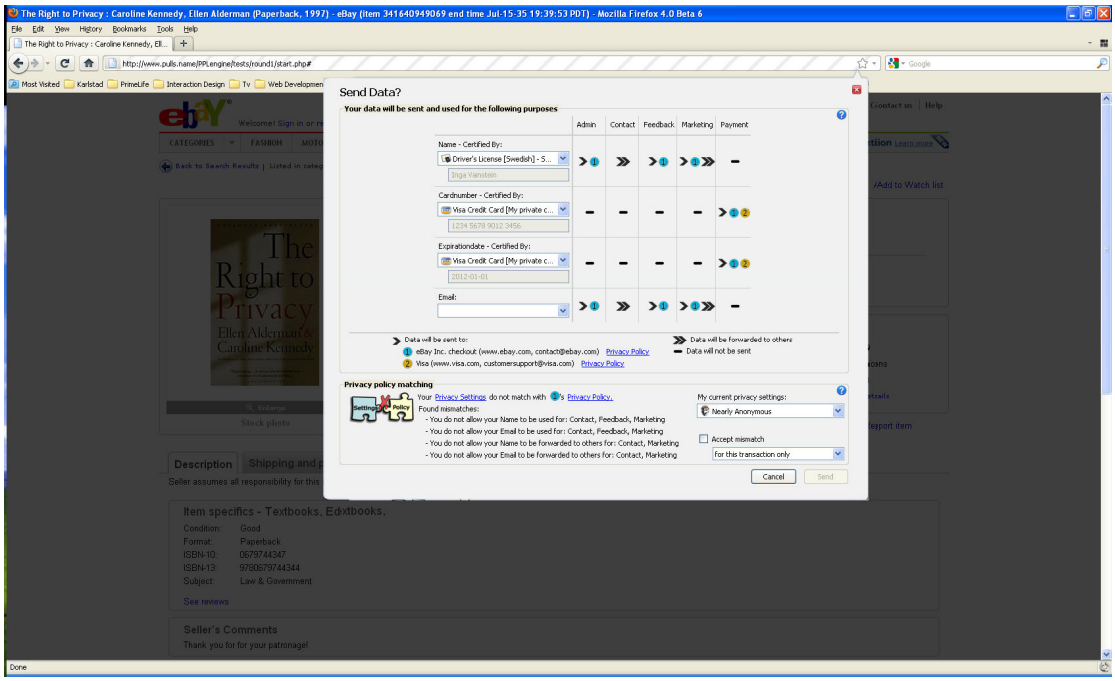

Figure 1: "Send Data?" dialog opens while the service provider's website is dimmed in the background

2.2 Warnings can cause rushed and unwanted user reactions

Common usability guidelines suggest expressing error messages in an understandable and consistent manner or, even better, to design and develop programs that minimize the use of disrupting warnings and annoying error messages by working properly [Nie90, CRD03]. In the case of security and privacy-friendly interfaces, it has been recognized that passive warnings or consistent error messages are often dismissed or not trusted by users [ECH08, V-SB10].

Some suggestions have been made in order to make users consciously aware of threats to their privacy and to promote informed decisions at the moment of disclosing personal private information. These suggestions often contradict known usability guidelines, which are not always applicable to the design of privacy-friendly interfaces. Egelman recommends interrupting the primary user task, providing choices on how to proceed safely, prevent habituating users to common warnings, and distorting the look of suspected websites so that users do not trust them immediately [ECH08]. Similarly, Villamarin-Salomon suggests the use of polymorphic warning dialogs to promote thoughtful responses to security dialogs [V-SB10].

Besides the risk that users might not notice the seriousness of warnings, from our experiences performing usability tests with PRIME and PrimeLife prototypes, we identified at least two more problems that privacy alerts must address:

- Users may try to get rid of intrusive privacy warnings by simply changing to less privacy-friendly settings, without being fully aware of what the consequences will be. In usability tests of early privacy policy management mock-ups, we experienced that very prominently displayed warnings, informing test users about a mismatch between her privacy preferences and the services side's privacy policy, led some test users to panic and prevented them from thinking through the consequences of their actions by changing to more "generous" privacy preference settings in order to eliminate the warning, instead of simply accepting the mismatch only for this case. Basically by this, they accepted implicitly not to be warned about

more privacy intrusive data handling practices for future transactions without that they were at that moment aware that this was the consequence.

- Our tests also showed that extensive warnings about a services side's privacy practices can be misleading for many users, because as discussed above, users often have problems differentiating between the user- and the services-side of an identity management system. Therefore extensive warnings can result in users losing their trust in the whole identity management system.

2.2.1.1.1 Lessons learned:

Privacy warnings should be carefully depicted by well chosen icons, colour code and text. Especially warning signs and red or yellow colour should be used with care and only in serious cases. During our work on privacy policy interfaces, we have been careful to provide users with feedback in relatively neutral manner about whether or not their privacy preferences are matched to the privacy policy of a service provider to prevent people from panicking. A rather discrete puzzle-piece icon has for instance been used to inform users about policy *matches* or *mismatches* in one of our policy interfaces, which has been successively improved within several iteration rounds. In the last test iteration of the PrimeLife policy interfaces, our 16 test subjects rated how easily these icons were understood with 4.1 on the average (where 5 was the highest and 1 the lowest value), with a standard deviation 1.15. No rushed user reactions were observed for this last UI iteration during the usability tests.

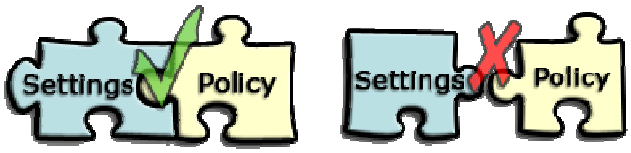

Figure 2: PrimeLife icons for displaying a match and a mismatch between the user's privacy preferences ("Settings") and the services side's privacy policy.

Usability tests have to evaluate carefully how warnings are understood and perceived by end users.

2.3 Trust in PETs

Shneiderman [Shn00] defines trust as the positive expectation that a person has for another person based on past performance and truthful guarantees. Trust in a system may be an important factor for acceptance of a system..

Already our usability tests of early PRIME prototypes, which we conducted in 2004, have shown that there are problems to make people trust the claims of a systems' privacy enhancing features [ACC+05]. Some participants voiced doubts over the whole idea of attempting to stay private on the Net. "Internet is insecure anyway" because people must get information even if it is not traceable by the identity management application, explained one test participant in a post-test interview. Another test subject stated: "It did not agree with my mental picture that I could buy a book anonymously". Also, in a study on the perception of user control with privacy-enhancing identity management solutions for RFID environments, test users lacked trust in proposed PET solutions, even though the test users considered the PETs in that study as fairly easy to use [GS05].

However, we observed that with the increased penetration of Internet usage and increased attention given to privacy topics and PETs, such as the anonymity service TOR, people have

gotten more familiar with the idea of PETs and this situation has probably changed slightly within the last years, but still remains a challenge.

In general, evaluating users' trust in a system is very difficult. During a usability evaluation of a system, users only have limited time familiarise themselves with the system and interact with it. They are "forced" to use this system, whereas in real-life they will only use it, when they are interested in the functionality of the system. Conclusions could be drawn implicitly, e.g. from questions such as "How much would you be willing to pay to use this system?" A user, who for instance answers "I would possible use a more developed version if it was free" was most likely hesitant to express his doubts of our prototype.

2.3.1.1.1 Lessons learned:

Trust in the claims that PETs provide certain privacy features, plays a key role in the acceptance and uptake of PET solutions. For novel PETs with a functionality, which may not fit the users' mental model of how the technology should work, users may however lack trust. The evaluation of users' trust in a system is very difficult and requires a careful design of post-test interview questions/questionnaires, which allow analysing indicators for (mis-)trust. Ideally, evaluations of users' trust in a PET would involve several factors, such as longitudinal studies and contextual usage in the users' daily lives.

2.4 Intercultural differences

Privacy is a cultural construct [LS93]. How privacy is defined and experienced can differ much between different cultures. Depending on the level of privacy protection in different countries, users in these respective countries may have different experiences, perception and knowledge of privacy concepts, e.g. of online privacy policies (which are required by law to be posted on European websites that are collecting personal data). This has in turn implications on how easily people from different cultures can understand user interfaces for illustrating these privacy concepts. Usability tests of privacy policy display and management user interfaces conducted at Karlstad University in the fall of 2010 with test students from Pakistan, Iran, Sweden and other European countries, showed for instance, that the concepts of privacy policy and privacy preferences were more difficult to understand for the Pakistani and Iranian students. Arguably, this can also be due to a different exposure to the Internet in those countries, and that their experiences buying products or paying for services over the Internet are not as high as they are for European citizens.

Within HCI research, much work has been done on intercultural interface design and the need for research of cross-cultural understanding of interface metaphors has been stressed [Eve98].

Within the scope of PrimeLife, we therefore also conducted an intercultural comparison test for the policy icons which we developed in WP4.3. The evaluation was conducted in the form of a paper mockup test with 17 Swedish and 17 Chinese students at Karlstad University in spring 2010. From the test results, it was obvious that the test subjects had different understandings of the icons because of their cultural backgrounds. While Swedish test persons had for instance no problems in understanding the "paragraph" for the purpose "legal obligations" or the "post horn" as an icon for the purpose "shipping" (see Figure 3), these icons were not understood by the

Chinese test subjects.

Figure 3: Example of policy icons, which were well understood by Swedish test students, but not understood by Chinese test students

2.4.1.1.1 Lessons learned:

Privacy concepts and user interfaces for illustrating them may be differently understood by individuals from different cultures.

Our intercultural icon comparison studies have demonstrated the fact that icons that are well understood in the western world are not necessarily easily understood by persons from another culture. When designing interfaces for privacy, icons and metaphors that fit the target population should be employed. If the interface is meant to be used by a variety of different cultures, then the icons and metaphors used should be understood by most people.

For user interfaces for PETs with a target group beyond European users, we recommend to conduct cross-cultural usability studies and intercultural usability comparison tests to test intercultural comprehension.

2.5 Comprehension of PET terminology

For the design of usable PET interfaces, it is very important that the employed wording is well understood by the users. Throughout several evaluations of the users' perception of privacy, security and PETs, as well as their understanding of related processes, have been investigated by CURE [GKW+10, GWKT10]. This research has underlined the need for understandable wording and further explanation of key terms. Terms used in PET interfaces should neither require a university degree in law nor in the field of security and privacy.

Research throughout the PrimeLife project has indicated that a quick evaluation of terms with only a few (non-expert) users can lead to indications on which words should be avoided in interfaces [KWGT09]. As part of the PrimeLife HCI activities, we have investigated several privacy terms that were also used in the PrimeLife prototypes [GWKT10].

We identified the following five terms as being easiest to understand:

- Privacy protection
- Required data
- Digital traces
- Identity management
- Full privacy policy

The terms rated as being very hard to understand are the following:

- Anonymous credentials
- Privacy preferences
- Linkability
- Privacy enhancing

2.5.1.1.1 Lessons learned:

The evaluation of the user's comprehension of privacy and PET-related terms can help to determine what terms should be avoided in PET user interfaces and replaced by alternative terms, which may be easier to comprehend by end users. PET terms to be used in UIs should therefore be evaluated for their comprehension by representatives of the target audience. For the PrimeLife policy interfaces, we chose, for instance, the technical term "privacy settings" instead of the term "privacy preferences" used in the PrimeLife Policy Language (PPL) specification, as the terminology evaluation conducted by CURE revealed that "privacy settings" was much better understood.

2.6 Data minimization is difficult to show

Data minimization is a fundamental privacy design principle which, in essence, requires that all applications and services should use only the minimal amount of data necessary for the transaction at hand. The objective is, of course, to preserve the privacy and minimize the possibility to profile users based on their behaviour. A key technology in achieving data minimization for applications are anonymous credentials [Cha85], [Bra99], [CL01]. A traditional electronic credential is a set of personal attributes that is bound to an individual by cryptographic means and that the user can use to prove these attributes. All usage of such a credential entails showing all attributes in the set, irrespective of the demands of the current transaction. In contrast, anonymous credentials let the user reveal any possible subset of attributes of the credential, characteristics of attributes, or prove possession of the credential without even revealing the credential itself. For example, a user with a governmentally issued anonymous driver's license credential, can, using zero knowledge proof, reveal and prove any one of the following; her birth date, her birth day, being over or under any given age, or the fact that she has a valid driver's license without revealing any other attributes of the credential.

In order to investigate the users' understanding of an anonymous credential selector GUI, we performed three rounds of tests based on different mental models of anonymous credentials [WF11]. The first round of tests was based on the card metaphor, i .e. users were asked to select the source cards of the credentials, whereas the second round of tests was based on an attribute based approach were the users were asked to select specific verified attributes they possessed. In the third round of tests we created a hybrid version of the two mental models building on the most positive results of both. In short, the results of the three rounds of tests regarding a selection mechanism for anonymous credentials show that the data minimisation properties are very difficult to show and that user's comprehension of the UIs clearly hinge on the induced mental model. Our results also show the detrimental effects of relying on analogies that only fit to a certain extent, since users are very stuck in their mental models. Furthermore we show that a possible remedy to this issue is to focus on the key difference between the old and the new technology (for example the adaptable card metaphor).

In addition to the main findings regarding the effects of mental models on users UI comprehension, there is also an interesting methodological implication of our work. Standard measurements during usability evaluations include task completion time and task success rate as measures of efficiency and learnability. Additionally, in order to assess the user value and satisfaction with the system, users are often asked if they enjoyed working with the system and if they would recommend the system to a friend. However, the validity of such data rests on the implicit assumption that task completion equals comprehension. In our studies, where comprehension was the focus, we explicitly investigated users' understandings of their actions.

Our results show that irrespective of level of understanding, we had a 100 percent success rate in terms task completion of the primary task. Our results also show that users liked our system and would recommend it to a friend, despite the fact that they did not understand the basic functionality of the system. Although there seemed to be problem in the users' understanding, they had the impression that the system was of use to them, otherwise they would not recommend it to third persons.

2.6.1.1.1 Lessons learned:

These results clearly show the need for including explicit measures of comprehension when testing the usability of complex tasks, especially if they are secondary tasks that rely on counterintuitive technologies such as anonymous credentials.

Our studies also showed that the mental models of users affect their understanding of the data minimisation property of anonymous credentials. A key issue for the wide use of privacy enhancing technology, such as anonymous credentials, is that, if users should understand and appreciate their privacy-enhancing features, or at least not misunderstand them, the mental model evoked by the user interfaces is of great importance. The question how the data minimisation property could be best conveyed to end users needs therefore especially be researched and investigated.

2.7 Conclusions: mental models are difficult with novel technology

As already pointed out in this section, one of the major obstacles in introducing a novel technology is describing it in such a fashion that the average user will comprehend the pros, cons, and benefits of the new system. The introduction of incremental innovations is most often easily framed in the terms of previous systems, i.e.," this system is faster or has more functionality built in than the predecessor". However, when it comes to radical innovations this is very difficult. Despite this, the most often used path to UI development are analogies to real world concepts or systems that the user already knows. The objective of using analogies is to help the user create a mental model of the system. A model that helps the users contextualise information in the interface or the system and aid the user in making predictions regarding the effects of various choices and actions. Failures in creating correct mental models of a system leave the user with an inadequate understanding of her actions.

The results of our user studies show that users often lack adequate mental models to protect their privacy. For instance, on a system level not understanding the flow of data in a network makes it impossible to anticipate sniffing, logging or man-in-the-middle attacks. On the same note, not understanding how an application routs your data makes it impossible to understand the privacy features of this application. On a user interface level, not understanding the meaning of icons such as the padlock makes it impossible to evaluate how secure a transaction is. Understanding the meaning of UI elements such as icons is obviously central to successfully using an application. Our user tests show that warning icons are an area where the effects of icons can be counter productive as users have a hard time discriminating between warnings regarding the privacy of the transaction and warnings regarding the configuration of the system. Thus, instead of evaluating the privacy levels of the transaction, users repeatedly reconfigured the application in order to '*solve the problem*' and make the warnings disappear and, in doing so, lower their level of privacy. Our work with a credential selection mechanism for anonymous credentials highlights the difficulties in using analogies when describing novel technology. In our first rounds of tests the majority of users believed that the anonymous credentials would work in the same fashion as the plastic credentials we compared them to. However, in our later test when we added reference to

the main difference between the two types of credentials (“adaptable”) and thus changed the induced mental model of the users error rates decreased by 64 per cent. A very clear finding in theses studies regarded the use of the Swedish personal number. As this number is widely used in Sweden, users anticipated that this number should be present in the transaction despite the fact that it was neither asked for nor shown anywhere in the interface.

These results all show that inducing adequate mental models is a key issue in successful deployment of novel privacy technology. When it comes to privacy, the effects of incorrect mental models lead to difficulties in using a given application or to not being able to take adequate steps in order to protect one’s information.

Chapter 3

How to design PET User Interfaces

Based on our own experiences, results from other researchers, and lessons learned from the PrimeLife project, this chapter provides some guidelines for the development and design process of user interfaces for privacy-enhancing technologies. First, we will provide HCI heuristics for PETs, which adapt, exemplify and extend classical HCI heuristics by taking PET-specific aspects into consideration. HCI heuristics are "rules of thumbs" which are usually complemented with best practice solutions and other HCI guidelines when designing systems. HCI patterns that we have developed in PrimeLife and that we briefly summarise in this chapter are merging such best practice solutions from HCI and different guidelines, including those that we derived from PrimeLife. They provide an approach to describe, organize und present solutions and best practices for design problems in the PET domain. Finally, we discuss the need for following a User-Centered Design (UCD) Process for developing PETs and show how HCI patterns and HCI heuristics for PET should be applied during the UCD cycles .

3.1 HCI Heuristics for PETs

The lessons learned from the PrimeLife project enable us to point out PET specific aspects of usability heuristics, to be considered in particular when conducting evaluations in the PET domain.

The main goal of these heuristics is to provide a means for IT professionals to conduct heuristic evaluations of their PET developments. Thus, they can, in particular, inspect privacy-enhancing technologies using traditional heuristics with a special focus on this selected application domain.

Based on years of usability engineering experiences as well as related research results of other usability experts, we adapted and exemplified the list of Usability Heuristics by Nielsen [Nie92, Nie94] and extended it with additional ones. The list of heuristics (which has been communicated to the consortium in deliverable "User Evaluation Plan" (H4.1.1)) is the following:

- **Consistency:** Consistency describes a common design of elements and processes from the users' point of view; all user interface concepts should thus be consistently designed
- **Feedback:** Feedback means that users expect a sufficient system reaction to all of their actions
- **Efficiency:** The user interface must enable the users to carry out their tasks efficiently

- **Flexibility:** The system must allow different users to work differently, or a single user to work differently if he wishes or needs to, in order to accomplish goals
- **Clearly marked exits:** The user must always know how he can leave a specific context, window or display when working with a user interface, and how he can return to his starting position
- **Wording in the user's language:** Wording in the user interface must be known and easily understandable to the user
- **Control:** The user must always be in control of the system; the user must never have the feeling of the system controlling him
- **Recovery and forgiveness:** The system must prevent the user from (unknowingly) taking severe actions; the user shall be able to undo changes or actions easily
- **Minimize memory load:** The user shall able to completely focus on his task, not being troubled with the user interface as such; therefore the user interface must require as little cognitive effort as possible
- **Transparency:** The user must always know what will happen when he takes an action- the user interface must be transparent
- **Aesthetics and emotional effect:** Everything has an emotional effect; if a user interface has an inappropriate emotional effect, it will interfere with the users' tasks

These recognized heuristics are applicable to the majority of interactive systems, and thus they can also be applied to the design of PETs, even though, according to our experiences and lessons learned, some adaptations are needed.

The following sections will highlight the impact of the PET-domain to each of these heuristics and represent them with examples and experiences from the PrimeLife project (Sections 3.1.1 – 3.1.11). Furthermore, PET-specific heuristics obtained through the research within PrimeLife will be presented in Sections 3.1.12 to 3.1.15. Thus, the PET-related aspects of traditional heuristics and the PET-related heuristics can be used, when conducting heuristic evaluations of PETs.

3.1.1 Consistency

Consistency is a well-known principle of design, which dictates that the look-and-feel, behaviours and actions of a product need to be uniform across the characteristic of the product. Some of the aspects of an interface that a designer usually looks for, when trying to add consistency to a product include the used wording, the interactions available from controls, the graphical elements, colours, warning messages, styles, etc.

From our experience in the PrimeLife project we have learned that for designing PETs the principle of consistency can be seen as twofold. On one hand, the interface of a PET should be consistent in its look-and-feel, especially when the user interacts with it for the purposes of manipulating it or *adjusting* it directly (i.e., modifying settings, changing passwords or keys, etc). For example, designing a privacy editor (Privacy Tuner) that can be accessed from within the "Send Data?" dialog (Figure 4), ought to have a similar look and interaction paradigm as the "Send Data?" dialog. Applying consistency in this case, would reduce the cognitive load of users and increase its usability.

On the other hand, designing a consistent look for error or warning messages can be counterproductive, resulting in users being numb to their effect and dismissing them without reading them [ECH08, V-SB10]. In this case the use of warning messages that look different every time they appear was suggested by [V-SB10] as a way of protecting the privacy of users by encouraging them to read the contents of the warning and making a more conscious decision.

PETs should be consistent in their look-and-feel, especially when the user interacts with them for the purposes of manipulating them or adjusting them directly. Warning messages should look different every time they appear to encourage users to read the content of the warning and make an informed decision.

3.1.2 Feedback

In an offline world, users are accustomed to get a reaction as a consequence to their action when interacting with physical objects. For example, when interacting with a light switch, the user gets an immediate response, or feedback, since the light bulb is turned on right away. In the same way, users expect to have an immediate reaction to their actions when interacting with technology. Feedback needs to be given fast and should be synchronized with the users' actions.

An example of the principle of feedback applied to PETs can be seen in the "Send Data?" dialog from the PrimeLife prototypes. When a user is carrying out an online transaction, the "Send Data?" dialog pops-up as soon as the service provider is requesting information from the user, that is, as soon as the user *clicks* on a "Submit" button. This immediate response to the user's actions by making the dialog appear as soon as the user clicks a button on a web service does not only provide users with feedback of what is going on, but also helps them protect their privacy.

PETs should help the users to be aware about privacy risks and provide feedback about the handling of their data and whenever their privacy is at risk.

3.1.3 Efficiency

Many new technologies have the purpose of enabling users to carry out some tasks efficiently, meaning that users should be able to increase their productivity with the help of the technology (note that it is the user's productivity that must be enhanced and not the computer's). Efficiency is not only measured by the time it takes for users to complete certain tasks, but also by the levels of the users' cognitive processes that the system forces them to activate.

Some sample guidelines to improve the users' efficiency when using a software system include, amongst many others, providing consistent dialog messages, using good defaults and structuring menu options logically. By adopting these guidelines, the users' cognitive processes are lowered, thus saving time and increasing efficiency.

PETs tend to be complicated systems for average users to understand. Efficient PETs should consider the well-known issue that "privacy" is rarely the primary concern of users who are trying to accomplished some other task [WT99]. The prototype for the Privacy Dashboard, for instance, is an example of how to empower users with information, while at the same time not interrupting the tasks they are trying to accomplish. The Privacy Dashboard displays a context-sensitive icon embedded in the web browser which informs users about the use of personal data of a visited website. If necessary or alarming, users are able to control the Privacy Dashboard via a few clicks, not disrupting them largely from their primary task and not affecting their productivity to a great extent, but helping them to protect their information and providing them with transparency.

PETs should consider that 'privacy' is a secondary concern of users and therefore need to empower them with information, while at the same time not interrupting the tasks they are trying to accomplish.

3.1.4 Flexibility

The *flexibility* heuristic states that the system must be flexible enough so that it adapts to the different needs of different users. For instance, expert users of a system may have different needs and requirements than beginners. Also, a user from a particular profession, for instance a doctor, might use spreadsheet software in a very different way than an accountant.

Flexibility is, to some extent, connected to the *efficiency* heuristic, in the sense that by providing interfaces that allow different users to work differently, they can increase their efficiency levels when trying to accomplish a goal.

PETs should also allow different ways of interaction, depending on the goals and needs of different users. In particular, PETs should satisfy the needs and curiosities of privacy concerned users, but should also be understandable and manageable by non-expert users while helping them protect their privacy. As an example from PrimeLife, the "Send Data?" dialog supports the needs of inexperienced users by providing them with standard privacy preferences which are easy to choose, depending on the transaction at hand. Users with higher, lower or more specific privacy concerns can customize these standard privacy levels to fit their own wishes for different situations easily "on the fly". Our usability tests showed the "on the fly" privacy preference management was well perceived by end users.

PETs should satisfy the needs and curiosities of privacy concerned users, but should also be understandable and manageable by non-expert users while helping them to protect their privacy.

3.1.5 Clearly marked exits

Users should always know how to navigate through the different contexts of a system. Furthermore, all parts of an application should be easily accessible by the users. Nevertheless, the users should have the possibility, at any time, to leave either the system or particular parts of the system in just a few clicks.

Examples from PrimeLife include the Privacy Dashboard and "Send Data?" prototypes. These prototypes pop-up as needed in order to protect the privacy of users. However, it is very simple for users to make them disappear and access them again as needed, without too much effort.

PETs should be designed in a way that they are not invasive, but always accessible and dismissible.

3.1.6 Wording in the users' language

A very important principle when designing complex user interfaces that are often grounded in technical concepts is to use of terminology and language that average users understand. At every step, users should be able to understand their options and possible consequences from their actions. In a sense, the understanding of the wording is the basis for a user's informed consent. If, for instance, terms displayed in privacy notices are misunderstood, users may agree to disclose their data, even though they would not have done so, if they had understood the privacy policy correctly.

In the context of PETs, the use of appropriate terminology for privacy is therefore of key importance. PrimeLife research has shown that users have trouble understanding privacy related

terms and the wording used in privacy policies of service providers [GWHW11]. For example, as discussed in Section 2.5, most users did not understand the terms “anonymous credentials”, “linkability” and “privacy enhancing”. Besides, it was found that most users understood and preferred the term “privacy settings” over the term “privacy preferences” [GWHW11].

The wording in PETs should be clear, simple, and understood by the majority of users. If necessary, usability studies should be conducted in order to find out if the terminology used in the PET is understandable by most users. Research throughout the PrimeLife project has indicated that a quick evaluation of terms with only a few (non-expert) users can lead to indications on which words should be avoided in interfaces [KWGT09]. Users should not require advanced knowledge of law, cryptography or other technical fields in order to benefit from a PET.

It should be also considered that the choice of words when displaying warning messages in PETs should be carefully thought-through, since users need to easily understand these messages to take appropriate actions.

Privacy terminology can be very complex and specific. It is important that the wording in PET user interfaces is clear, simple, and understood by the majority of users.

3.1.7 Control

The control heuristic dictates that the user must always be in control of a computer system, instead of feeling that the system is the one in control. A.S. Patrick et al. define control in the context of PETs as “the ability of the user to perform some behaviour… users must be aware of the need to act before they can execute the behaviour” [PKHvB03]. The property of transparency, presented later, serves as one of the prerequisites for imposing users’ control over their personal data. One way of making the user feel in control of the system is by keeping them informed about what the system is doing by providing appropriate feedback within reasonable time.

Making users feel in control of their personal data is an important requirement for the development of PETs. Endowing the users with control cannot only increase the users’ trust in the system, but it can also help users make appropriate informed decisions when managing their personal data.

The prototypes for the Privacy Dashboard, the Data Track and the “Send Data?” dialog, are all examples of PETs that provide transparency and give users control of the system and over their own personal data. These prototypes allow users to see where their data will or has been sent to and the purposes for which it is used. Moreover, the Privacy Dashboard and the Data Track allow users to correct or delete the personal data that they have submitted earlier, i.e., giving the users control over their previously disclosed data. Also, the “Send Data?” dialog lets users control different levels of privacy preferences by managing and customizing the existing levels.

Users have to feel in control of their personal data at any time when using PETs.

3.1.8 Recovery and forgiveness

A recommended guideline for most interactive systems is that it should prevent the users from consciously performing unwanted actions, and in case unwanted actions are done, the system should give users the possibility to undo them. A common example of this heuristic is the “undo” button in many software applications.

The issue of recovery and forgiveness is a very critical issue for PETs. Once personal information is disclosed or compromised on the Internet it is often almost impossible for users to recover that

data and undo the damage. Currently users possess a mental model suggesting that once their data is submitted to a service provider they have lost control over it.

Therefore technologies that want to protect the users' privacy must take into consideration the prevention of errors.

As mentioned earlier, PrimeLife's Data Track prototype is an example of a PET that gives users the capability to delete or correct their data located on the services' side and thus provides an "undo" option.

PETs should prevent users from compromising personal data (error prevention). In case information has been disclosed accidentally, users should be supported in recovering.

3.1.9 Minimize memory load

System designers should try to minimize the users' memory load by increasing the visibility of interactive elements, accommodating affordances, and supporting intuitive interactions. In general, the system should not force users to remember information, but instead should aid them to remember previously learned information. The purpose is to let users focus on the task at hand, instead of increasing their cognitive load.

An example of this heuristic can be seen in the "Send Data?" prototype, which is divided into two main panels. The top panel displays a summary of the service provider's privacy policy in the form of a two-dimensional table showing the data being requested and the purposes for which it will be used. The bottom panel gives users information on whether their privacy settings match or do not match the service provider's privacy policy (summarized in the table above).

Usability tests of the "Send Data?" dialog revealed that users that have their privacy settings memorised looked whether their privacy settings *mismatched* the privacy policy by looking first at the two-dimensional table. However, the "Send Data?" dialog reduces the users' memory load by displaying a visual representation of the mismatch in the bottom panel, and does not force the user to remember his own privacy settings. Nevertheless, the use of too many visual cues can also overwhelm users, thus a balance must be achieved between helping the user recognize screen elements while keeping the interface relatively simple and clean.

PETs should not increase the users' cognitive load, for example, by requiring the user to adopt new and unused interaction paradigms.

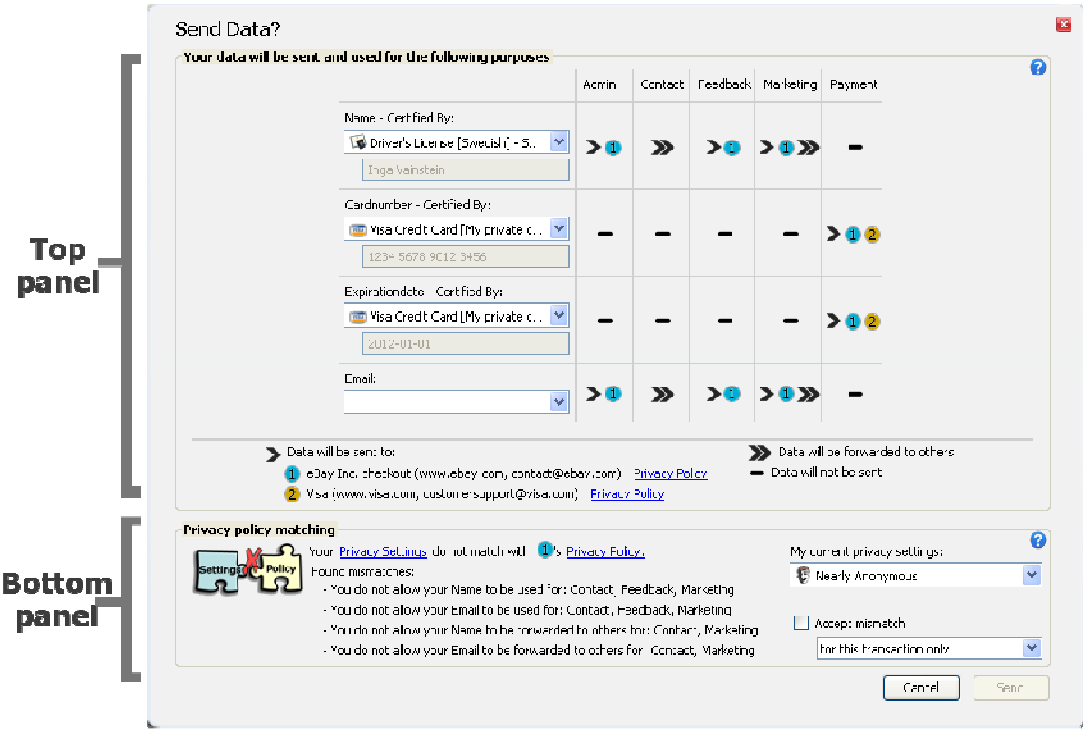

Figure 4. The "Send Data" dialog design aims at minimising the users' memory load

3.1.10 Transparency

Transparency refers to the property of letting users know what goes on *behind the scenes* as they interact with a digital artefact. By endowing technology with the property of transparency, it becomes easier for users to understand how the technology is constructed, how it works and how information is processed, which can lead to a better understanding of the implications, scope and ways to adapt or change the technology [LS04].

For PETs, transparency is a key property to consider from the beginning of their design. By letting users know how the PET works, and how it processes their personal information and what is happening behind the scenes, the users' trust can be gained and maintained. Some of the prototypes done for the PrimeLife project, such as the "Send Data?" dialog, the Data Track or Privacy Dashboard, carry the property of transparency by letting the users know how their information will be or has been processed on the internet, or how a data controller is planning to use their information.

PETs should provide information to the users about how they work, how they process their personal information and what happens behind the scenes. All processes have to be transparent and understandable to the user.

3.1.11 Aesthetics and emotional effect

As every human made artefact has aesthetic and emotional effects this also relates to PETs. In the PrimeLife project we learned that the aesthetics of the design has an impact on the potential use of PETs.

During the project we saw other influences of the aesthetics in the PET domain. The aesthetical quality of the visually perceived UIs is reduced by the presentation of encrypted text. Encrypted text is an aesthetic disruptor to the end-user.

By applying the heuristic of "aesthetics and emotional effect" we propose a solution such as well designed icons (an aesthetic icon) instead of the aesthetic disruptor. These icons have to be understandable by the user and therefore need a well planned strategic design process involving users.

PET design has to follow aesthetic principles and underline its main purpose.

The sections above presented the original heuristics adapted and exemplify to the usability of PET interfaces. The following sections present PET-specific heuristics identified through the research carried out within PrimeLife.

3.1.12 Remote vs. local handling of data

With the introduction of the Internet, users became accustomed to send their personal data to different data controllers to get the benefits of a service. Moreover, nowadays users are storing their data on remote servers that allow them to access these data from wherever they are at any time. Privacy policy and transparency tools try to protect the users' personal data by making them consciously aware which data are going to be released, the consequences of releasing their data, which data has already been released and the steps they can take to protect such data.

However, as discussed in Section 2.1, this increasing flow of data from the user to the service's side and vice versa can become very complicated and counterintuitive for users who are unaware of the actual location of certain pieces of data. In other words, at times, many users do not know whether the data they manage and the decisions they take will have local or remote effects. Therefore, it is very important for a PET interface to make a clear distinction between the data being handled remotely on the server's side and the data handled locally on the user's side.

PET interfaces should make it clear to the user and help him understand what data is being handled remotely on the service's side and what data is handled locally on the client's side.

3.1.13 Internationalization

As mentioned in Section 2.4, the concept of privacy and perceptions of privacy risks can vary from one culture to another. Studies within PrimeLife have shown that icon images are interpreted or understood differently by Swedish and Chinese test subjects [HNH11]. However, not only icons might be culturally-dependent, but also the privacy terminologies or the used wording, the awareness of privacy risks online, the consequences of disclosure of private data, the level of exposure and concerns to the invasion of privacy, and other factors.

The design of PETs should keep in mind that privacy is a culturally formed construct if global solutions for protecting privacy across individuals from different cultures need to be provided.

Interfaces for PETs should consider the intercultural aspect of privacy and privacy risks.

3.1.14 Informed consent

An essential aspect of some PETs is that they can help users make informed decisions about their actions. Consent has been identified by A.S. Patrick et al. as an important HCI requirement for PET interfaces [PKHvB03]. Consent refers to the explicit, conscious, awared and informed decisions of users to agree to the way their personal information is going to be handled by a data controller. Just In Time Click-Through Agreements (JITCTAs) have been suggested as a way to

enforced users' informed consent to the disclosure and processing of personal data [PKHvB03]. In case information is very sensitive, the use of double-JITCTAs has also been proposed.

A.S. Patrick et al. suggest the following points in order to comply with the informed consent prerequisites imposed by the European Privacy Directive:

- "give informed consent to the processing of [Personal Data];
- give explicit consent for a Controller to perform the services being contracted for;
- give specific, unambiguous consent to the processing of sensitive data;
- give special consent when information will not be editable;
- give consent to the automatic collection and processing of information" [PKHvB03].

Besides JITCTAS, user interface concepts for supporting users to provide well informed consent include multiple-layered privacy notices suggested by the Article 29 Working Party [Art04], Drag-and-Drop-Agreements (DaDAs) and the "Send Data?"-dialog developed in the PRIME and PrimeLife projects [Pri415] [PFHD+05].

Interfaces for PETs should enable users to make informed decisions about their personal data disclosure.

3.1.15 Good privacy-friendly defaults

The inclusion of good defaults in computer programs has already been recognized as an important pattern for the design of interfaces in general [Tid05]. Providing good defaults become especially important for PETs, since most users seldom want to be bothered about configuring the program to protect their privacy. Even if they wanted to, many users are not aware about the best ways of protecting their privacy. It is therefore crucial that PET interfaces, do not only provide the controls necessary to allow users configure their privacy preferences in a friendly way, but also contain good privacy-friendly or privacy-promoting default settings, options and behaviours.

The "Send Data?" dialog, for example, has controls defining three default levels of privacy that the user can adjust depending on a particular transaction. More experienced or concerned users are also allowed to personalize these levels of privacy by creating custom values. The interface saves these custom values so that users can access them in future transactions.

Similarly, the default behaviour for the Clique prototype is to let users create different faces and group contacts into collections with different privacy settings. By providing this default behaviour, Clique assists users at protecting their privacy in social network sites, which is the opposite approach from other social networks such as Facebook, which default settings are set to display most information to all contacts.

Standard configurations and settings should be provided to users that are privacy friendly by default.

3.1.16 PET Usability Checklist

The above mentioned HCI Heuritics for PETs have been transformed into a PET Usability checklist. The appendix holds 25 points to be considered when designing PETs. It is meant as decision support tool for decision makers and technology builders.

3.2 HCI Patterns for PETs

In addition to HCI Heuristics, HCI Patterns are an important instrument for guiding the UI design. Patterns are a useful approach to describe, organize und present solutions and best practices for design problems, which are based on long-term experiences. Although much work can be found concerning either patterns or privacy, work focusing on patterns for Privacy Enhancing Technologies (PET) is very rare.

Within the PrimeLife, we have developed HCI Patterns for PETs, which are part of Deliverable D4.1.3 [Pri413] and was also presented in [GWG10]. The developed UI patterns shall help designers and developers of PETs creating useable and understandable interfaces for end-users.

Our Pattern approach contains fifteen patterns, including patterns dealing with the display of privacy policies, privacy icons and policy icons, privacy awareness panel in collaborative workspaces, informed consent, secure passwords, privacy-aware wording, credential selection, trust evaluation of services sides, Data Track, Privacy Options in Social Networks, Selective Access Control in Forum Software, Privacy Enhanced Group Scheduling.

A crucial factor, when developing patterns is that they must be consistent with other patterns in the collection. The reason for this need for consistency is that patterns describe not only a solution for a special problem, but a solution for a special problem in a particular domain [KMP08] and thus, should not contradict other patterns in the same domain. This means that all patterns must direct towards the same purpose – in our case, all of our HCI PET-patterns have to support the creation of user interfaces for PETs.

During the development of the patterns we combined guidelines and already proven approaches from the field of HCI. We also integrated the knowledge, experience and results, which we gathered in the PRIME and PrimeLife projects into our PET Patterns.

The goal of the developed patterns is to present complex technical PET mechanisms in an understandable way for users to help designers and developers to create usable and supportive interfaces for PETs.

Through various end-user tests we were able to identify problems of different patterns and were therefore able to fix them in later versions of the patterns. This process is visible through a 5 star-rating, which we used (where 0 stars mean that no end-user tests were conducted and 5-stars mean that much end user testing was done and the results prove the content of the pattern). The whole pattern collection can be found on the PrimeLife website (available at [Pri413]).

3.3 User-centric design for PETs

Figure 5 shows the user centred design process (UCD) as defined in “ISO/TR 16982:2002: Usability methods supporting human-centred design”.

When the need for a UCD is identified, the ISO model provides four main activities:

1. Specify the context of use: this identifies the context the users and user-groups will use the system in.
2. Specify requirements: Based on user-goals the requirements are defined.
3. Create design solutions: This part of the process may be done in stages, building from a rough concept to a complete design.
4. Evaluate designs: Usability evaluations with real users should be conducted to evaluate the designs from the previous step.

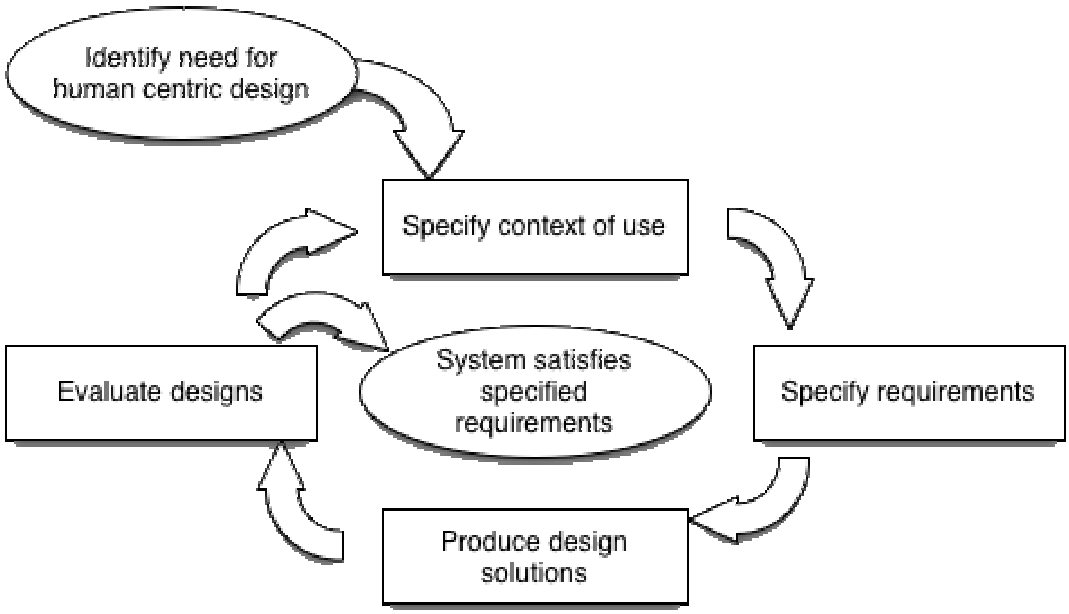

Figure 5 The UCD based on ISO/TR 16982:2002

The process is iterative: The steps in the diagram are likely to be repeated in an iterative development process until the system satisfies the specified requirements. Hence, the design solutions get more and more user-centric during such iterations.

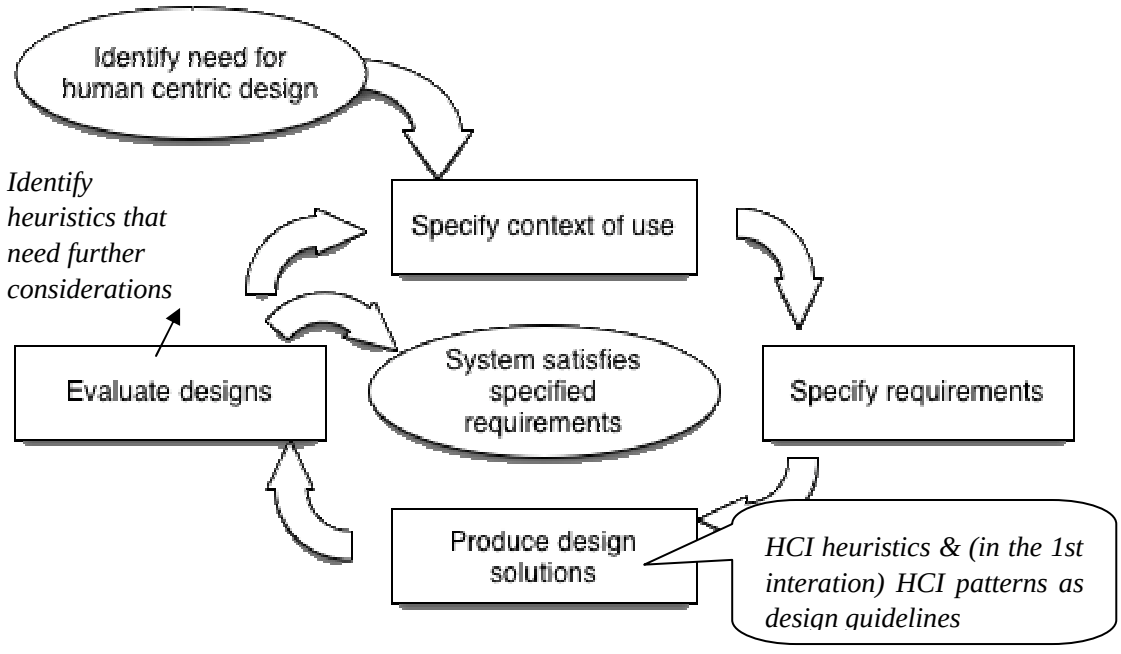

Figure 6 Applying HCI heuristics and HCI patterns during the UCD cycles

In Figure 6 we show how our HCI heuristics and HCI patterns that we developed for PETs can be applied in the design and evaluation phases of the UCD cycles.

In PrimeLife we saw, that at least 4 iterations are needed when designing UIs for PETs (which tends to be more than for standard software where experience says that about 3 iterations are what happens in practice). The iterations are fed with different HCI input:

- Iteration 1: When producing the design solutions for the first iteration we used the patterns as a basis (D4.1.3 HCI Pattern Collection). In addition, the HCI heuristics are applied (see Section 3.1). As mentioned above, HCI heuristics, which are rules of thumb, and HCI patterns, which are best practice solutions, complement each other. For instance, an HCI pattern for the display of privacy policy is based on a multi-layered design of privacy policies as recommended by the Art. 29 Working Party [Art04]. Such a multi-layered design should in addition fulfil rules of consistency, feedback, efficiency, etc. When evaluating the design in the first iteration we saw that an expert based heuristic evaluation (see Chapter 3.1) found potential for improvement, which was the basis for a re-design for the second iteration. Hence, in this evaluation phase as well as in the evaluations of all later iterations, the HCI heuristics that need further consideration in the next re-design are identified.
- Iteration 2: The design solutions of the second iteration can again be checked against the heuristics (see Chapter 3.1) – especially those which have not been successfully met in the previously tested design of Iteration 1. For the second iteration we recommend to already include end-users in the evaluations. We learned that at this stage more informal usability tests with end-users are efficient.
- Iteration 3: Based on the outcome of iteration 2 the re-design should be done. After that we recommend a full usability-laboratory test. We saw that – even when theory suggests that around 10 users per test are sufficient [Nie92] – it pays off in the PET domain to go up to 16 participants per test.
- Iteration 4: After the findings of the usability-laboratory tests have been considered we have seen that from this point we can show prototypes to the world and start a public beta testing phase.

As the ISO standard does not specify exact methods we highlight the UCD subtleties we have been applying within the PrimeLife project.

During the course of development of the many PrimeLife prototypes implemented by different PrimeLife partners, we had the chance of being involved at the various stages of the development processes. In these processes, some of the PrimeLife development teams considered HCI aspects from the beginning, others realized the need of HCI input during their development, and yet some others attempted to apply HCI principles on top of their already developed prototypes. At the end, this gave us a unique opportunity to compare the results of the different prototypes, with the added variable of the point in time, at which HCI practices were considered during their development.

We present our observations in this section, making the point that the implementation of PETs, as most other interactive systems, should consider the end users from the beginning, consult known HCI guidelines and consider the advice from experts during the whole development process. In other words, the development of PETs should go through the various stages suggested by the UCD process as depicted in Figure 5.

Table 1 presents the different PrimeLife prototypes in relation to their average level of satisfaction, as reported by participants of the usability tests performed at CURE, and the time that usability evaluations and the opinion from HCI experts from PrimeLife Activity 4 were introduced into the development of the prototype (early in the development process, sometime in the middle of the development process, or late, when the implementation was practically finished).

Table 1. PrimeLife prototypes showing their average satisfaction rating and the time when PrimeLife HCI activity was involved in the development process

Prototype	Average Satisfaction Rating (1=very high, 5=very low)	Standard Derivation	Approximate time when HCI expert opinions were considered (Early, Middle, Late)
"Send Data?" Dialog	1.83	0.37	Early
Privacy Dashboard	1.85	1.02	Early
Dudle	2.25	0.69	Middle
Wiki	2.31	0.75	Middle
Clique	2.71	0.82	Late
Scramble!	Not ready for testing because HCI issues were recommended to be fixed first	-	Late

We want to convey that following a user-centric process of design, where the involvement of HCI guidelines, principles and opinions of usability experts is of great importance for the development of usable PETs, provide higher levels of user satisfaction. Table 1 reveals a relation between the level of satisfaction and the time in the development when the HCI expert opinion and usability tests from PrimeLife Activity 4 were taken into consideration. Note, however, that this table only shows that there may be a tendency of higher user satisfaction when there is higher input from HCI activities, but that no statistically significant conclusions can be made from this table. The low number of test participants (n=16) and to the fact that the prototypes have different functionalities that appeal to different people can be variables that affect the tendency shown in the table. Therefore, no definite conclusions can be drawn here.

When designing for other common interactive systems, development teams have the chance to apply good enough "HCI guesses", which are usually based on real-world metaphors, previous user experiences interacting with similar technologies, or well-grounded mental models. Whitten at al. already showed that good enough "HCI guesses" may not work for designing usable security and privacy [WT99]. Also it could be argued that in contrast to other interactive systems, the use of HCI guesses for the development of complex PET systems will most likely tend to result in products with low levels of user satisfaction, as seen from the tendency shown in Table 1. It is therefore recommended to involve usability experts and interaction designers early in the process of PET development, who can provide input to the context of use of the PET, gather requirements for the users' ultimate needs and tasks, propose usable design solutions for PETs and carry out usability evaluations.

The reason for early involvement of HCI experts in case of PETs is that users often have problems articulating their privacy needs, since privacy is often a secondary priority when they are trying to achieve a goal. Therefore, gathering requirements by simply asking users to list a series of things a PET system should include in order to help them protect their privacy, is not an appropriate method for eliciting requirements for PETs. Similarly, letting developers imagine what the needs of the users are might be not appropriate, since the mental models of system developers and the

mental models of users are rarely the same. Dedicated usability experts could help gathering the requirements from users by employing different sociological and HCI methods, as well as studying the users' mental models with regards to privacy. As an example from the PrimeLife prototype evaluations, we can see in Table 1 that the "Send Data?" dialog interface, which was subject to a user-centric design process and various iterative cycles, in which usability testing was performed during every cycle, had high ratings of user satisfaction. On the other hand, the prototype for Scramble!, for which the PrimeLife HCI Activity was only involved after its deployment (although the developers themselves performed short usability tests with high-school students), did to our knowledge in principle not follow a user-centric design approach, was not considered ready for usability testing due to problems of comprehensibility of its UIs, even by other privacy and security experts, which needed to be fixed first.

Chapter 4

How to evaluate PET User Interfaces

As pointed out in Chapter 3, people often do not seem to have correct mental models of privacy and security technologies [KCJ09, FHH02], and privacy and security are not the primary goals of users [WT99]. Thus, the planning and execution of usability tests for PETs has to be done carefully and meticulously, keeping in mind these arguments.

Our experience in testing prototypes developed for PrimeLife has proven that certain factors can have an influence on the outcome of PET usability tests. The following sections illustrate some of these factors and provide guidance that should be taken into account when designing and carrying out usability tests for PETs.

4.1 Things to consider before the evaluation

The preparation before a usability test is a fundamental part of the test. It usually involves writing a test plan, setting up a testing environment, preparing the required test material and recruiting test participants. In the following sections we present some of the factors that we found to be important in the process of planning for a usability test based on our experiences executing usability tests for PETs for PrimeLife.

4.1.1 Recruiting participants

The participants who attended the numerous usability evaluations during the PrimeLife project at CURE and KAU were recruited in different ways. While CURE had a large database of participants, KAU recruited participants mainly from the university campus.

The main benefit of the database approach is that recruiting participants with different demographic backgrounds (e.g., level of education, age, etc.) can be done easily by selecting a specific subsample from the participants stored in the database. In this way, it was possible to compare the level of understanding of participants with different backgrounds and the problems they had with the tested prototypes. Nevertheless, with the database approach great amounts of time might be spent on inviting participants for an evaluation. Participants have to be contacted by telephone, and setting a strict time schedule for the evaluations is necessary, since participants usually have a defined time slot in which they can come to participate on the test.

Recruiting participants from the University campus has the advantage that tests can be carried out almost 'spontaneously', since participants (usually students) were readily accessible and willing to participate. This opportunity allowed for an easier iterative approach to user interface design and faster rounds of testing due to the availability of participants. However, this approach had the disadvantage that the results might mainly reflect the opinion and interaction behavior of students and university staff, which limits the findings to a constricted sample and might not be generalized to represent the opinions of a wider population. On the other hand, recruiting participants on campus had the additional benefit that responses can be obtained from participants with various cultural and educational backgrounds. Nevertheless, as mentioned in Section 2.4, during the usability evaluations at KAU, it was noticed that participants' responses to privacy related questions were somewhat dependent on these cultural differences. Therefore, the recruitment of participants for evaluating PETs ought to consider that cultural differences in privacy constructs exist.

Also, from our experience in PrimeLife we noticed that eliciting responses from student participants enrolled in more technical fields, such as computer security, cybercrime, networking or other subjects alike, could produce more accurate responses in regards to their beliefs on how their personal information is handled online and the protection of their privacy. Evidently, students with such technical interests have often already developed the right mental models of the PET technology being tested, or are able to evoke the comprehensive mental models based on previous experiences with these kinds of technologies. Carrying out usability tests with this type of participants might returned biased responses that do not reflect the opinion of users who are not computer-savvy. On the other hand, students or staff coming from other educational backgrounds do not have the same level of understanding about technologies that help them protect their safety and privacy online, and using them as test participants could provide better insights on the needs and thoughts of normal users. However, it all comes down on the kind of people that the PET is targeted for at the end.

When recruiting participants for a PET usability evaluation, consider that privacy is a cultural construct. Participants with technical knowledge might already have mental models that fit the technology. Consider the pros and cons of the different recruiting approaches.

4.1.2 Adapting to the users' language

When planning a usability test for PETs, it is important to use terminology that participants will understand. Tasks should be presented in an understandable way so that participants are able to fulfil them. This also means that, when designing tasks for a usability evaluation of a PET, the target group for which the prototype is designed should be kept in mind, and evaluators should remember that the thing being tested is the usability of the PET prototype and not the users' knowledge of privacy related terms, vocabulary and concepts. For example, trying to explain the term "partial identity" would not make sense to most test participants; instead it should be described in more detail, such as "a partial identity is an identity which represents a special part of your life (sport club, family stuff, and so on)." More information about the lessons learned on the appropriate wording for PETs can be found on Section 2.5.

At the same time, the participants' native language should be used as much as possible. In different international projects it became obvious for CURE that even German speaking participants who stated that they have good knowledge of English had problems understanding the wording and labelling of user interfaces presented in this language. It is therefore very hard to differentiate between general usability problems of the software program and problems occurring due to the participant's misunderstandings of words, labelling and terminologies. Especially when evaluating privacy software, the used labels and terms should be chosen carefully. Participants should have the possibility to ask the test evaluator for a translation or a clarification.

During PrimeLife's usability evaluations, CURE also provided all material for the tests such as questionnaires and tasks in the German language. The reason was that participants were not English native speakers and therefore it was easier for them to read and understand German tasks and executing them even when the user interface was presented in English. Furthermore special attention was paid to the wording of the tasks, for example, when creating tasks for Clique special attention was placed in not translating words and labels which were essentials for users in order to complete a certain task, such as the term "Face".

Computer security and privacy terminologies are not intuitive for ordinary users and non-native language speakers. Careful use of words and terms needs to be considered at each step of the design of a usability evaluation.

4.1.3 Designing tasks for testing privacy concepts

As mentioned before, security and privacy are only secondary goals of the user [WT99]. In other words, users want to accomplish tasks (e.g., a shopping transaction) and do not want to be bothered configuring their privacy settings or reading privacy warnings. It could be argued that a PET with good usability remains almost invisible, but still guards the user's privacy. At the same time, there are occasions in which a PET should make the users' consciously aware that their information is being requested, disclosed and/or compromised. These opposite use-cases create a challenge when designing test scenarios for evaluating usability and user acceptance of PETs.

Often, when evaluating the usability of PETs, designing tasks and scenarios that are not directly related to privacy issues, but that rise possible privacy concerns on the user's periphery, should be carefully considered. For instance, during the evaluation of the "Send Data?" dialog [AFPK11], an e-Shopping scenario was presented, in which the participants' main goal was to buy a product on eBay. The shopping transaction requested the participant to submit private information, and the dialog was shown as soon as information was requested by the service provider (cf. JITCA [PFHD+05] [PKHvB03]). In this scenario, the user was made aware that information was requested when the dialog did pop-up. In this case, the usability of this PET should not only measure the user-friendliness of its interface, but also how it helped users to accomplish their goal, how secure it made users feel, how much the user appreciated that their privacy was being protected or how disrupted the users felt by the appearance of the PET dialog.

On the other hand, testing the usability of the social network Clique required, for example, an almost 'transparent' and subtle method. When interacting with a social networking site, users usually have the aim of connecting with other people and sharing their momentary live experiences with others. Testing for privacy repercussions, e.g. when users decide to share personal information with the world, while at the same time, they wish to retain a certain level of privacy, is not easy, and test scenarios should be designed carefully and iterative, so that their validity and feasibility is accurate.

One important factor, which should be considered when designing tasks for usability evaluations, is that users are aware that it is just a test and not a real-life scenario. It could be argued that users react in a different way when using fake trial data for the evaluation than when using their own real personal data. Nevertheless, it is possible to test the general understanding of end-users concerning the evaluated prototype even if their behaviour is not exactly the same compared to a real-life situation.

For evaluating the users' understanding of the functionality of a PET, is it necessary to provide meaningful scenarios and well-designed tasks to the participants. These scenarios and tasks should consider that privacy is a secondary goal at times, but other times attention from the user is required.

4.2 Things to consider during the evaluation

After having looked at the most important factors to be considered before the actual execution of a PET usability evaluation, we now present some of the factors we identified as important at the moment of performing the actual evaluation with the recruited test participants.

4.2.1 Introduction to the test - creation of mental models

One of the main challenges when evaluating PETs is that many of the concepts introduced by these technologies are usually unknown to the test participants, since they do not possess the appropriate mental models.

A mental model is an explanation of a thought process about how something works in the real world. It is an explanation of a person's perceptions about what will happen when they act or react in a certain way. Our mental models shape our behaviour, including how we approach tasks. Mental models are the root reasons why a person does something in a particular way; they are built over a long time of experience and are, therefore, resilient. Consequently, a mental model provides a deep understanding of people's motivations and thought processes [Joh86, Jon95, You08]. Therefore, it is of crucial importance to use appropriate methods that carefully present the purpose of the PET being evaluated in understandable terms to the test participants in order to elicit the correct mental model for the PET.

During the evaluation of various PrimeLife prototypes, different methods were used to introduce participants to the tests. For example, one of the iterations of the credential selection tests used video prototyping to present the concept of anonymous credentials to participants. Evaluations of the "Send Data?" dialog prototype used an e-shopping scenario familiar to most participants and an introductory text explaining the steps of the test.

In contrast to other common software, participants rarely have a mature mental model of online privacy and information flow; therefore, creating the appropriate mental model consistently across participants of a PET evaluation becomes of crucial importance to obtain consistent results. Giving a different description and introduction to different participants of the same test might have an impact on the mental models they create and therefore influence their responses.

Rubin & Chisnell also support the idea that during usability testing the introduction to a test should be done in a consistent manner, in order to minimize influencing the participants' perception of the product [RC08]. It was corroborated during the evaluations of the PrimeLife prototypes that the wording of texts and other introductory material must be done with special care and in a consistent way. For instance, during the most recent evaluation of the credential selection concept, the word "adaptable card" was inserted into the test description. Participants were told that an "adaptable card" would be generated that would only adapt the necessary personal information requested during an online transaction. Results indicated that, by introducing different wording participants created a different mental model as compared to those from previous tests.

Participants rarely have a mature mental model of online privacy and information flow. The choice of methods used to introduce the participant to the PET during the evaluation can influence the mental model they create. Introduce participants to the test in a consistent and structured manner.

4.2.2 Collection of demographic information

When conducting usability evaluations it is very helpful to collect demographic data about the participants. This data could include age, gender, cultural background and level of education. Other questions depend on the context of the evaluation. In PrimeLife, CURE additionally asked for the average internet usage and the use of a web browser. Further questions dealt with privacy awareness (e.g., "Are you concerned about your privacy on Web?"), registration in a web shop or community platform and frequency of reading privacy policies.

Collecting such a variety of information allows searching for significant differences between various groups of users (e.g., elderly people vs. younger ones, or expert vs. novice users). This can lead to implicit suggestions for design improvements tailored to certain user groups. However, when drawing conclusions from a demographic group of people a bigger sample is desired, since it can provide more statistically significant results. In this case, the suggestion from a user-centric design approach of testing fewer participants at every iteration cycle might not give enough evidence to make general conclusion about a whole demographic area.

Collecting demographic data about the participants could prove valuable at the moment of analyzing the test results.

4.2.3 Performing the usability evaluation

During the evaluations carried out at CURE, participants were asked to interact with the PET for a period of 2 – 3 minutes until they became more familiar with it. At Karlstad University, tests were designed so that the first task constituted a way to familiarise the participants with the prototype and the basic purpose behind it. Participants are also asked to think aloud while interacting with the tool, which allows the test evaluator to uncover usability and comprehension problems.

Another factor to consider during design and evaluation of PETs is that, in some cases, there must be a balance between convenience for the users and the protection of their privacy [FGSB08]. Furthermore, it should be kept in mind that different tasks involve different privacy risks. In other words, when evaluating the usability of PETs, the activities that users are performing and the contexts in which they are performed, are essential in order to estimate the severity of the privacy risks.

As discussed in section 2.6, our evaluation of the credential selection mechanisms for anonymous credentials revealed problems of inducing mental models for novel technologies by means of analogies. Even though an analogy might sound feasible at first and it helps the user build a rudimentary understanding of the system, the performance of the users has to be thoroughly investigated. In the case of PETs, it is especially important to evaluate users' understanding of the consequences of their actions, as they might complete the tasks without any problem, but at the same time without understanding the implications of their actions. Our solution to this problem has been to use an iterative process, where we have evaluated a UI, both in terms of task completion and in terms of users understating (see Section 4.2.4). Based on these results we have redesigned the UI and the mental model we have tried to induce to accomplish both, an easier interaction flow and a higher level of understanding.

During the evaluation of a PET consider all the factors that might have an influence on the participants' perception of the privacy risks, such as the context of use, the balance between privacy and convenience, etc.

4.2.4 Using post-questionnaires for evaluating PETs

Since privacy is usually only a secondary task for users, user interactions for managing privacy should be minimised. Thus, the use of post-questionnaires can be valuable at revealing the opinions and experiences of the test participants after using the PET. The main purpose of using post-questionnaires during usability testing is to gather information about the participants' opinions and feelings, to clarify their earlier responses and to uncover further confusions that they might have had during the test [RC08]. Important things to find out include the appreciation of the participants on how the PET helped them to protect their privacy, how well the interface was understood, how obtrusive or supportive it was at accomplishing a task, the extent to which they are willing to configure it, learn it and use it.

During the various evaluations at CURE, the following questions were usually asked to participants after they interacted with the PET:

- How satisfied were you with the software? (5 Item Likert Scale)
- Would you recommend the software to a friend? (Yes – no)
- Which was your general impression of the software? Please motivate your answer.
- Was there anything you did not understand? Please motivate your answer.
- Was there anything you missed while using the software? Please motivate your answer.
- Can you imagine using such software? Why? Why not?
- How much time will you spend at configuring / learning the software?
- Do you have any further suggestions, wishes or comments?

These questions were targeted towards technology acceptance and their user experience of the entire system. They allowed us to get further insights into users' understanding of the concepts and their mental attitudes about the evaluated software and therefore let us derive the users' opinions concerning PETs. The answers provide insight into errors or problems which occurred during the evaluation and may not be notices from the test-supervisor. Furthermore these questions allow us to draw conclusions on users' general opinion about the software. Especially when they won't recommend the software to a friend we can conclude that either the usability was not well or the purpose of the software was not so obvious.

Questions that can be used to reveal the constructed mental model of the test participants can also be asked at the time of eliciting the participants' judgment of the PET during the completion of each tasks or during a post-test interview. Such questions include:

- Was your data well protected when you tried to achieve *a task*?
- What information did others find out about you when you tried to achieve *a task*?
- Who has access to your information after you completed *the task*?
- How secure do you feel after having completed *the task?*
- What was your experience while accomplishing *the task*?

In addition to the post-test questions suggested above, we also developed and used PET-USES, described in the following section, which can provide more objective and measurable results with regards to the technical functionality of the PET tested.

Since interaction with a PET is minimal at times, using post-questionnaire is a valuable way to obtain the participants' experiences and opinions about the PET, as well as their level of understanding of the purpose of the PET.

4.2.5 PET-USES

Usability evaluations of PETs are, in many ways, not different from any other usability tests. However, in examining the usability of PETs it is important to also investigate the users' understanding of the application and its usage after the users have had a chance to interact with the given prototype.

In a number of our user tests during PrimeLife we have noticed that users might very well solve a given task satisfactorily and subsequently say that they liked the application and would recommend it, but, when asked about the consequences of their actions it turns out that they have not understood the main point of using the application. The problem is that the current questionnaires for measuring user experience, usability and various HCI (Human-Computer Interaction) aspects such as the hedonic quality [Has03] of both, software and websites [Bro96][TS04] focus on the usability of the primary task of the system.

The PET-USES (Privacy-Enhancing Technology Users' Self-Estimation Scale) [WWK09] is a questionnaire that enables users to evaluate PET-User Interfaces both in terms of the primary task and specific PET related secondary tasks. Thus, the PET-usability scales have a dual purpose. They evaluate the system's general usability and the extent to which the system assists the user in learning and understanding privacy related issues. In combination with measuring the explicit comprehension of the underlying technology of the system the PET-USES adds a lot of knowledge to a usability evaluation.

As with all questionnaires, in order to get valid data, successfully using the PET-USES requires larger tests samples than the ordinary fast iterative design cycle permits. It is thus not recommended to use the PET-USES to evaluate small changes in one design iteration. However, collecting data throughout the design process and only comparing bigger blocks comprised of several iterations makes it possible to use the PET-USES in combination with fast iterative design cycles. PET-USES is a valuable supplement to other HCI questionnaires for evaluating interfaces and a good approach to investigate the participants' understanding of the consequences of their actions when using a PET or involve themselves in a privacy-related activity online.

4.3 Conclusion

In this Chapter we have presented factors that we deemed important when planning and performing usability evaluations of privacy-enhancing technologies. These factors are based on our experience in executing usability evaluations for PrimeLife project prototypes, and include the following ones:

- When recruiting participants, aspects such as their cultural and technical background need to be considered
- The wording and privacy terminology used in the evaluations needs to be carefully chosen, so that they are understood by all participants

- Test tasks need to be designed with care for evaluating the user's understanding of the PET functionality
- Introducing participants to the tests in a consistent and comprehensible manner
- Collection of demographic data that is valuable for later analysis in particular regard to the test user's technical and cultural backgrounds
- Considerations at the moment of carrying out the evaluation in regard to factors that might have an influence on the participants' perception of privacy risks
- Use of post test questionnaires and PET-USES as valuable tools for obtaining a more accurate account of the experience and opinions of the test participants.

Chapter 5

Conclusions

This deliverable should serve as an experience report from the PrimeLife project, which allows other developers and designers that plan to develop user interfaces of privacy-enhancing technologies, to learn about special HCI challenges and typical HCI fallacies, which especially arise in the PET domain and that need to be considered. It therefore also provides guidance on how the design and evaluation of PET user interfaces can address these issues. This deliverable should thus help UI developers to avoid doing typical mistakes and provides at the same time HCI heuristics, best practice solutions and guidance for the development of usable PETs.

Several of the HCI challenges discussed also necessitate further research for enhancing the usability of specific PET solutions. In particular, we have pointed out that inducing adequate mental models for novel PET technologies, which are unfamiliar to non-technical users and for which no good analogies exist, remains a challenge of key importance for making their interfaces usable.

Appendix: PET Usability Checklist

1. Are you consistent with current PETs (terminology, icons, concepts)?
2. Are all parts and patterns of the PET-UI consistently designed?
3. Do you mutate warning messages in a way they appear different every time?
4. Are your feedback messages to the end-users persuasive enough to be read?
5. Is the information given sufficient enough to enable users to conduct an informed decision?
6. Do you make users aware of privacy risks?
7. Do you provide sufficient feedback about the handling of user-data?
8. Are users warned when their private data is at risk?
9. Don't you interrupt the users' task with unnecessary hurdles in their workflow?
10. Do you describe technical terms in a way that is understandable to the end-users?
11. Is sufficient background information available for curious users?
12. Is there a possibility for end-users to dismiss intervening PETs?
13. Is the PET designed in a way that it does not hinder the end-user to accomplish their primary goal?
14. Are the privacy related benefits of your PET clear to the user?
15. Are the used terms in the interface clear, simple, and understood by the majority of users?
16. Can the user control which data is disclosed or not?
17. Is the user forced to disclose more data than needed for the task at hand?
18. Can users correct unwanted disclosure of their data?
19. Does the UI prevent the user from compromising personal data?
20. Can the user apply simple, natural interaction models (instead of learning new concepts)?
21. Do you force the user to adopt to new and unfamiliar interaction paradigms?
22. Does the UI give the end-user a glimpse on how the background mechanics schematically work?
23. Are means implemented which support the user to understand what happens behind the scenes?
24. Does the aesthetical appearance of the PET fulfil the aesthetical expectations of the user (does it look "modern" instead of "old fashioned")?
25. Are the default settings privacy friendly?

References

[ACC+05] Andersson, C., Camenisch, C., Crane, S., Fischer-Hübner, S., Leenes, S., Pearsson, S., Petterson, J., Sommer, D., “Trust in PRIME”, Proceedings of the 5th IEEE Int. Symposium on Signal Processing and IT, December 18-21, 2005, Athens, Greece.

[AFPK11] Angulo, J., Fischer-Hübner, S., Pulls, T. & König, U. To appear in 2011, "HCI for Policy Display and Administration" in PrimeLife - Privacy and Identity Management for Life in Europe, eds. J. Camenisch, S. Fischer-Hübner & K. Rannenberg, Springer, pp. 261.

[Art04] Article 29 Working Party. 2004, Opinion on More Harmonised Information Provisions 1198704/EN WP 100, European Commission.

[BMG01] Blakley, B., McDermott, E., Geer, D. 2001.“ Information security is information risk management.” In Proceedings of the 2001 Workshop on New Security Paradigms ACM Press, New York, NY, pp. 97-104.

[Bra99] Brands, S. Rethinking Public Key Infrastructure and Digital certificates – Building in Privacy. PhD thesis. Eindhoven. Institute of Technology. 1999.

[Bro96] Brooke J. SUS: a "quick and dirty" usability scale. In: Jordan PW, Thomas B, B. A. Weerdmeester, McClelland AL, editors. Usability Evaluation in Industry London; 1996.

[Cha85] Chaum, D. Security without Identification: Transaction systems to make Big Brother Obsolete. Communications of the ACM, 28(19): 1030-0144, October 1985.

[CL01] Camenisch, J. and Lysyanskaya, A. Efficient non-transferable anonymous multi-show credential system with optional anonymity revocation. In Advances in Cryptology - Eurocrypt 2001, volume 2045, pages 93–118, 2001.

[CRD03] Cooper, A., Reimann, R. & Dubberly, H. (2003). *About Face 2.0: The Essentials of Interaction Design.* (1st edn.). New York, NY, USA: John Wiley \& Sons, Inc.

[ECH08] Egelman, S., Cranor, L.F. & Hong, J. (2008). You've been warned: an empirical study of the effectiveness of web browser phishing warnings. In *Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems.* New York, NY, USA: ACM. 1065.

[EU95] European Parliament, 1995. Directive 95/46/EC of the European Parliament; available at: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:31995L0046:EN:HTML

[Eve98] Evers, V. 1998. Cross-cultural understanding of metaphors in interface design. *Ess, C.and Sudweeks, F., Proceedings CATAC,* 98

[FGSB08] Franz, E., Groba, C., Springer, T. & Bergmann, M. 2008. A Comprehensive Approach for Context-dependent Privacy Management. In *The Third International Conference on Availability, Reliability and Security.* IEEE. 903.

[FHH02] Friedman, B., Hurley, D., Howe, D. C., Felten, E., and Nissenbaum, H. 2002. “Users' conceptions of Web security: a comparative study.” In CHI '02 Extended Abstracts on Human Factors in Computing Systems CHI '02. ACM, New York, NY, pp.746-747.

[GWG10] Graf, C., Wolkerstorfer, P., Geven, A., Tscheligi, M. 2010: A “Pattern Collection for Privacy Enhancing Technology”, Proceedings of the Second International Conferences on Pervasive Patterns and Applications, PATTERNS ’10, Lisboa, Portugal.

[GWHW11] Graf, C., Wolkerstorfer, P., Hochleitner, C., Wästlund, E. & Tscheligi, M., to appear in 2011, "HCI for PrimeLife Prototypes" in PrimeLife - Privacy and Identity Management for Life in Europe, eds. J. Camenisch, S. Fischer-Hübner & K. Rannenberg, Springer, pp. 217.

[Has03] Hassenzahl, M., 2003. The thing and I: understanding the relationship between user and product. In: Blythe, M., Overbeeke, C., Monk, A.F., Wright, P.C. (Eds.), Funology: From Usability to Enjoyment. Kluwer, Dordrecht, pp. 31–42.

[HNH11] Holtz, L.-E., Nocun, K., Hansen, M. (eds.): Towards displaying privacy information with icons. Proccedings of IFIP/PrimeLife SummerSchool 2010, Springer Verlag (to appear 2011)

[Jac08] Jacob, F. 2008. Ästhetik und UX: Das Potential von Serious Motion Graphics, Xtopia 2008

[Joh86] Johnson-Laird, P. N. 1986. Mental Models: Towards a Cognitive Science of Language, Inference, and Consciousness. Harvard University Press, Cambridge, MA, USA.

[Jon95] Jonassen, D.H. 1995. Operationalizing mental models: strategies for assessing mental models to support meaningful learning and design-supportive learning environments. In The first international conference on Computer support for collaborative learning (CSCL '95), John L. Schnase and Edward L. Cunnius (Eds.). L. Erlbaum Associates Inc., Hillsdale, NJ, USA, 182-186.

[KCJ09] Klasnja, P., Consolvo, S., Jung, J., Greenstein, B. M., LeGrand, L., Powledge, P., and Wetherall, D. 2009. "When I am on Wi-Fi, I am fearless: privacy concerns & practices in everyday Wi-Fi use. " In Proceedings of the 27th international Conference on Human Factors in Computing Systems (CHI '09.ACM, New York, NY, pp. 1993-2002.

[KWGT09] Christina K¨offel, Peter Wolkerstorfer, Arjan Geven, and Manfred Tscheligi. A study on dynamic vs. static display of privacy preferences, 2009.

[KMP08] Khazanchi, D., Murphy, J., Petter S. 2008. Guidelines for evaluating patterns in the IS domain. MWAIS 2008 Proceedings, p. Paper 24. http://aisel.aisnet.org/mwais2008/24

[LS93] Lundheim, R., and Sindre, G. (1993). Privacy and Computing: A Cultural Perspective. In: Proceedings of the IFIP WG 9.6 conference on *Security and Control of Information Technology in Society*, edited by Richard Sizer te al., North-Holland, 1993.

[LS04] Löwgren, J. & Stolterman, E. (2004). Thoughtful interaction design: A design perspective on information technology. The MIT Press.

[Nie92] Nielsen, J. 1992. Finding usability problems through heuristic evaluation. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (CHI '92), Penny Bauersfeld, John Bennett, and Gene Lynch (Eds.). ACM, New York, NY, USA, 373-380. DOI=10.1145/142750.142834 http://doi.acm.org/10.1145/142750.142834

[Nie94] Nielsen, J., Heuristic evaluation. In Nielsen, J., and Mack, R.L. (Eds.), Usability Inspection Methods, John Wiley & Sons, New York, NY, 1994.

[GS05] Günther O, Spiekermann S. RFID and the perception of control: The consumer's view. Communications of the ACM 2005;48(9):73-6.

[PFHD+05] J.S. Pettersson, S. Fischer-Hübner, N. Danielsson, J. Nilsson, M. Bergmann, S. Clauss, T. Kriegelstein, and H. Krasemann. Making PRIME Usable. In SOUPS 2005. Symposium on Usable Privacy and Security, Carnegie Mellon University, 2005.

[PKHvB03] A.S. Patrick, S. Kenny, C. Holmes, and M. van Breukelen. Human computer interaction. In J.J. Borking & J.G.E. Olk G.W. van Blarkom, editor, Handbook of Privacy and Privacy-Enhancing Technologies: The Case of Intelligent Software Agents, pg. 249–290. College Bescherming Persoonsgegevens, Den Haag, The Netherlands, 2003.

[Pri123] PrimeLife WP1.2. PrimeLife Privacy Dashboard. In D. Raggett, editors, PrimeLife Deliverable D1.2.3. PrimeLife, http://www.primelife.eu/results/documents, 2011.

[Pri413] PrimeLife WP4.1. HCI Pattern Collection. Deliverable D4.1.3. PrimeLife, http://www.primelife.eu/results/documents, 2010.

[Pri414] PrimeLife WP4.1. High-level Prototypes. In C. Graf, P. Wolkerstorfer, E. Wästlund, (eds.), PrimeLife Deliverable D4.1.4. PrimeLife (http://www.primelife.eu/results/documents). 2010

[Pri415] PrimeLife WP4.1. Final HCI Report. In Simone Fischer-Hübner et al., editors, PrimeLife Deliverable D4.1.5. PrimeLife, http://www.primelife.eu/results/documents, 2011.

[Pri422] PrimeLife WP4.2. End User Transparency Tools: UI Prototypes. In Erik Wästlund and Simone Fischer-Hübner, editors, PrimeLife Deliverable D4.2.2. PrimeLife, http://www.primelife.eu/results/documents, June 2010.

[Pri432] PrimeLife WP4.3. UI Prototypes: Policy Administration and Presentation – Version 2. In S. Fischer-Hübner and H. Zwingelberg, editors, PrimeLife Deliverable D4.3.2. PrimeLife, http://www.primelife.eu/results/documents, June 2010.

[RC08] Rubin, J. & Chisnell, D. 2008. *Handbook of Usability Testing: How to plan, design and conduct effective tests.* (2nd edn.). Wiley-India.

[SF05] Sasse, M.A., Flechais, I. 2005. Usable security: What is it? how do we get it? InLorrie Faith Cranor and Simson Garfinkel, editors, Security and Usability: Designing Secure Systems that People can Use. O'Reilly Books, 2005.

[Shn00] Shneiderman, B., 2000. Designing trust into online experiences. Communications of the ACM, 43: 57-59.

[TS04] Tullis, S. & Stetson, J. A Comparison of Questionnaires for Assessing Website Usability. Usability Professional Association Conference; 2004.

[Tid05] Tidwell, J. & Ebooks Corporation. 2005. Designing interfaces: Patterns for effective interaction design. O'Reilly.

[V-SB10] Villamarín-Salomón, R.M. & Brustoloni, J.C. 2010. Using reinforcement to strengthen users' secure behaviors. In *Proceedings of the 28th international conference on Human factors in computing systems.* ACM. 363.

[Was10] Wash, R. 2010. "Folk models of home computer security." Symposium on Usable Privacy and Security (SOUPS), July 14-16 2010, Redmond WA, USA

[Wel08] Welie M.v. 2008.: Patterns in Interaction Design. Available at: http://www.welie.com/patterns/index.php

[WT99] A.Whitten and J. D. Tygar. Why Johnny can't encrypt: A Usability Evaluation of PGP 5.0. In SSYM'99: Proceedings of the 8th conference on USENIX Security Symposium, pages 14–14, Berkeley, CA, USA, 1999. USENIX Association.

[WWK09] Wästlund, E., Wolkerstorfer, P., Köffel, C., 2009. "PET-USES: Privacy-Enhancing Technology - User's Self-Estimation Scale," Pre-Proceedings of Fifth IFIP/PrimeLife Summer School, Nice, France, 2009.

[WF11] Wästlund, E., Fischer-Hübner, S. The Users' Mental Models' Effect on their Comprehension of Anonymous Credentials, in PrimeLife - Privacy and Identity Management for Life in Europe, eds. J. Camenisch, S. Fischer-Hübner & K. Rannenberg, Springer, pp. 233.

[You08] Young, I. 2008. “Mental Models: Aligning Design strategy with human behavior.” Rosenfeld Media, New York.